

USDA PRIVACY IMPACT ASSESSMENT FORM

Project Name: Strategic Data and Analysis (Data Mining/Data Warehouse)

Description of Your Program/Project:

To create a single warehouse of crop insurance data. This data relevant data mining and statistical tools to conduct research to decrease program vulnerability; detect fraud, waste, and abuse created by insured producer, agents, and loss adjusters. Enhance integrity of FCIC in compliance with 7 USC 1514 § 515(i)(2) of the Federal Crop Insurance Act.

DATA IN THE SYSTEM

1. Generally describe the information to be used in the system in each of the following categories: Customer, Employee, and Other.	Customer data that is used by the system is farm and producer data. As well as agent and adjuster data.
2a. What are the sources of the information in the system?	Economic data, NIASS data, soils data, M-13 data (1991-2003) through the Data Acceptance System, USDA drought data, NOAA Weather Data, RMA Compliance Tracking System, GIS data
2b. What USDA files and databases are used? What is the source agency?	The information is stripped from RMA's Policy Holder Database that is populated by the Data Acceptance System.
2c. What Federal Agencies are providing data for use in the system?	USDA – RMA; NOAA
2d. What State and Local Agencies are providing data for use in the system?	N/A
2e. From what other third party sources will data be collected?	N/A
2f. What information will be collected from the customer/employee?	N/A

3a. How will data collected from sources other than the USDA records and the customer be verified for accuracy?	N/A
3b. How will data be checked for completeness?	Data is checked upon entry into the databases by automated and manual methods.

ACCESS TO THE DATA

1. Who will have access to the data in the system (Users, Managers, System Administrators, Developers, Other)?	Only the DBAs will have direct access to the data. This data is reformatted from the DAS data into one that is easier to use for data mining. The data is then read into the database/warehouse where it is stored read-only. Developers can query the data, but they will only see results from their queries. The users will be able to view the results from the data, but will not be able make queries or view the raw data.
2. How is access to the data by a user determined? Are criteria, procedures, controls, and responsibilities regarding access documented?	The user contacts the Security Liaison Representative (Susan Hughes) who is also the COR for the system. The SLR will complete the FCIC-586 user request form and send them to the RMA Security Staff for final approval.
3. Will users have access to all data on the system or will the user's access be restricted? Explain.	Each user will only be granted the minimum access required to perform their job.
4. What controls are in place to prevent the misuse (e.g. browsing, unauthorized use) of data by those having access?	Each function performed on the system is explicitly audited and is entered into a project form to track their progress. Management reviews each query before it is allowed in production.
5a. Do other systems share data or have access to data in this system? If yes, explain.	Yes. Magnum will pull data results from this system.

5b. Who will be responsible for protecting the privacy rights of the customers and employees affected by the interface.	Privacy rights will be protected by the security controls implemented by the System Administrators
6a. Will other agencies share data or have access to data in this system (International, Federal, State, Local, Other)?	No
6b. How will the data be used by the agency?	The data will be used to check for patterns in crop insurance filings to highlight areas of fraud, waste, and abuse.
6c. Who is responsible for assuring proper use of the data?	The proper use of the data will be determined by the "customer" of this project, in this case RMA.

ATTRIBUTES OF THE DATA

1. Is the use of the data both relevant and necessary to the purpose for which the system is being designed?	Yes, the data used was specifically tagged for this purpose.
2a. Will the system derive new data or create previously unavailable data about an individual through aggregation from the information collected?	No. Data will not be altered or otherwise generated.
2b. Will the new data be placed in the individual's record (customer or employee)?	No. However, data may be used for further investigation by RMA.
2c. Can the system make determinations about customers or employees that would not be possible without the new data?	Yes. The sole function of the system is to present patterns in ways that were very difficult to impossible for a person to come up with.
2d. How will the new data be verified for relevance and accuracy?	The data is stripped from the Policy Holder Database. This database is populated by DAS, which has the function of verifying data.

3a. If data is being consolidated, what controls are in place to protect the data from unauthorized access or use?	N/A.
3b. If processes are being consolidated, are the proper controls remaining in place to protect the data and prevent unauthorized access? Explain.	N/A
4a. How will the data be retrieved? Can it be retrieved by personal identifier? If yes, explain.	Data is retrieved through the use of generated queries. Each query is labeled according to the developer. This label is a function of user name. The system's end users can view data that is delivered to them via a password protected user interface, or via email.
4b. What are the potential effects on the due process rights of customers and employees of: <ul style="list-style-type: none"> • consolidation and linkage of files and systems; • derivation of data • accelerated information processing and decision making; • use of new technologies. 	Due process could be affected by random and arbitrary searches that could raise suspicion to illegal activities.
4c. How are the effects to be mitigated?	Queries are only for investigative purposes. The queries are a support for further investigation, but due not in themselves provide for legal action.

MAINTENANCE OF ADMINISTRATIVE CONTROLS

1a. Explain how the system and its use will ensure equitable treatment of customers and employees.	Customers are not singled out by this system. Queries are based on groups and then are pared down by processes to ensure the greatest probability of matches.
2a. If the system is operated in more than one site, how will consistent use of the system and data be maintained in all sites?	N/A

2b. Explain any possibility of disparate treatment of individuals or groups.	The system is designed to detect fraud, waste, and abuse. Its sole purpose is to detect patterns by groups or individuals.
2c. What are the retention periods of data in this system?	Data is retained by the system for at least 10 years; this allows it to detect broad patterns.
2d. What are the procedures for eliminating the data at the end of the retention period? Where are the procedures documented?	No procedures have been developed as of yet, the system is designed to retain data for an indefinite amount of time.
2e. While the data is retained in the system, what are the requirements for determining if the data is still sufficiently accurate, relevant, timely, and complete to ensure fairness in making determinations?	Data is loaded monthly from RMA. This data will update all data on the system.
3a. Is the system using technologies in ways that the USDA has not previously employed (e.g. Caller-ID)?	Yes. The data-mining project is new to RMA.
3b. How does the use of this technology affect customer/employee privacy?	The data has a minimal impact on privacy, since this data is maintained at both the Insurance Providers and at RMA.
4a. Will this system provide the capability to identify, locate, and monitor <u>individuals</u> ? If yes, explain.	Yes. The system is designed to track fraud, waste, and abuse of the Federal Crop Insurance Program by producers, agents, and loss adjusters.
4b. Will this system provide the capability to identify, locate, and monitor <u>groups of people</u> ? If yes, explain.	Yes. The system is designed to track fraud, waste, and abuse of the Federal Crop Insurance Program by producers, agents, and loss adjusters.
4c. What controls will be used to prevent unauthorized monitoring?	No real-time monitoring occurs. Explicit requirements for queries have been set up in the contract. In addition, queries are developed upon request of RMA.

5a. Under which Systems of Record notice (SOR) does the system operate? Provide number and name.	FCIC-10 "Policyholder"
5b. If the system is being modified, will the SOR require amendment or revision? Explain.	N/A